American Journal of Medical Genetics (Neuropsychiatric Genetics) 67:580–594 (1996)

Rapid Publication

Additional Support for Schizophrenia Linkage on Chromosomes 6 and 8: A Multicenter Study

Schizophrenia Linkage Collaborative Group for Chromosomes 3, 6 and 8*

In response to reported schizophrenia linkage findings on chromosomes 3, 6 and 8, fourteen research groups genotyped 14 microsatellite markers in an unbiased, collaborative (New) sample of 403-567 informative pedigrees per marker, and in the Original sample which produced each finding (the Johns Hopkins University sample of 46-52 informative pedigrees for chromosomes 3 and 8, and the Medical College of Virginia sample of 156-191 informative pedigrees for chromosome 6). Primary planned analyses (New sample) were two-point heterogeneity lod score (lod2) tests (dominant and recessive affected-only models), and multipoint affected sibling pair (ASP) analysis, with a narrow diagnostic model (DSM-IIIR schizophrenia and schizoaffective disorders). Regions with positive results were also analyzed in the Original and Combined samples. There was no evidence for linkage on chromosome 3. For chromosome 6, ASP maximum lod scores (MLS) were 2.19 (New sample, nominal p = .001) and 2.68 (Combined sample, p=.0004). For chromosome 8, maximum lod2 scores (tests of linkage with heterogeneity) were 2.22 (New sample, p=.0014) and 3.06 (Combined sample, p = .00018). Results are interpreted as inconclusive but suggestive of linkage in the latter two regions. We discuss possible reasons for failing to achieve a conclusive result in this large sample. Design issues and limitations of this type of collaborative study are discussed, and it is concluded that multicenter follow-up linkage studies of complex disorders can help to direct research efforts toward promising regions.

KEY WORDS: schizophrenia, genetic linkage, collaboration, polymorphism, genotype

*The Schizophrenia Linkage Collaborative Group for Chromosomes 3, 6 and 8, coordinated by Douglas F. Levinson, comprises the following research groups (in alphabetical order) and co-authors:

Bonn/Haar/Jerusalem (Bonn) Dieter B. Wildenauer, Sibylle G. Schwab, Margot Albus, Joachim Hallmayer, Bernard Lerer, Wolfgang Maier *Department of Psychiatry, University of Bonn (D.B.W., S.G.S., W.M.); State Mental Hospital, Haar, Germany (M.A.); Graylands/UWA Clinical Research Unit, University of Western Australia, Perth (J.H.); Department of Psychiatry, Hadassah-Hebrew University Medical Center, Jerusalem (B.L.)*

University of Edinburgh Douglas Blackwood , Walter Muir, David St Clair, Stewart Morris *MRC Human Genetics Unit, Western General Hospital and University Department of Psychiatry, The Royal Edinburgh Hospital*

Iceland/Germany Hans W. Moises, Liu Yang, Helgi Kristbjarnarson, Tomas Helgason, Claudia Wiese *Department of Psychiatry, Kiel University Hospital (H.W.M., L.Y., C.W.); Department of Psychiatry, National University Hospital, Reykjavik (H.K., T.H.)*

(Author list continued on next page)

Received for publication September 13, 1996.

Address reprint requests to Douglas F. Levinson, M.D, AUH-EPPI, 3200 Henry Avenue, Philadelphia, PA 19129

Institute of Psychiatry/Cardiff (IOP/Cardiff) **David A. Collier, Peter Holmans, Jo Daniels, Mark Rees, Philip Asherson, Queta Roberts, Alastair Cardno, Maria J. Arranz, Homero Vallada, David Ball, Hiroshi Kunugi, Robin M. Murray, John F. Powell, Sin Nanko, Pak Sham, Mike Gill, Peter McGuffin, Michael J. Owen** *The Institute of Psychiatry, Departments of Psychological Medicine and Neuropathology, Section of Molecular Genetics, London (D.A.C., M.J.A., H.V., D.B., H.K., R.M., J.F.P., P.S.); Departments of Psychological Medicine and Medical Genetics, University of Wales College of Medicine, Cardiff (P.H., J.D., M.R., P.A., Q.R., A.C., P.M., M.J.O.); Department of Psychiatry, Teikyo University School of Medicine, Tokyo (S.N.); Department of Psychiatry, Trinity Centre for Health Sciences, St. James's Hospital, Dublin (M.G.)*

Johns Hopkins University Collaboration (JHU Collab) **Ann E. Pulver, Stylianos E. Antonarakis, Robert Babb, Jean-Louis Blouin, Nicola DeMarchi, Beth Dombroski, David Housman, Maria Karayiorgou, Jurg Ott, Laura Kasch, Haig Kazazian, Virginia K. Lasseter, Erika Loetscher, Hermann Luebbert, Gerald Nestadt, Carl Ton, Paula S. Wolyniec** *Department of Psychiatry and Behavioral Sciences (A.E.P., V.K.L., G.N., P.W.) and Department of Pediatrics (R.B., L.K.), Johns Hopkins University, Baltimore; Department of Genetics, University of Geneva Medical School (S.E.A., J.-L. B.); Instituto Di Psychiatria, Seconda Universita Degli Studi De Napoli, Naples (N.D); Department of Genetics, University of Pennsylvania Medical Center, Philadelphia (B.D., H.K.); Center for Cancer Research, Massachusetts Institute of Technology (D.H., C.T.); Human Neurogenetics Laboratory (M.K.) and Laboratory of Statistical Genetics (J.O.), The Rockefeller University, New York; Sandoz Pharma, Ltd. (H.L.)*

LGN-CNRS **Claudine Laurent, Michel de Chaldee, Florence Thibaut, Maurice Jay, Daniele Samolyk, Michel Petit, Dominique Campion, Jacques Mallet** *Laboratoire de Génétique Moléculaire de la Neurotransmission et des Processus Neurodégénératif, Centre National de la Récherche Scientifique, Hôpital de la Pitié Salpêtrière, Paris (C.L., M.d.C., D.S., J.M.); CHS Sotteville les Rouen (F.T., D.C.); CHS Saint Paul, La Reunion, France (M.J.)*

Medical College of Virginia (MCV) **Richard E. Straub, Charles J. MacLean, Stephen M. Easter, F. Anthony O'Neill, Dermot Walsh, Kenneth S. Kendler** *Department of Psychiatry (R.E.S., C.J.M., S.M.E., K.S.K.) and Department of Human Genetics (C.J.M., K.S.K.), Medical College of Virginia, Virginia Commonwealth University, Richmond; Department of Psychiatry, Queens University, Belfast (A.O.); The Health Research Board, Dublin (D.W.)*

National Institute of Mental Health (NIMH) **Pablo V. Gejman, Qiuhe Cao, Elliot Gershon, Judith Badner, Ethiopia Beshah, Jing Zhang** *Clinical Neurogenetics Branch, National Institute of Mental Health (USA)*

St. Mary's Hospital Medical **School Brien P. Riley, Swarnageetha Rajagopalan, Mpala Mogudi-Carter, Trefor Jenkins, Robert Williamson** *Deparment of Biochemistry and Molecular Genetics, St. Mary's Hospital Medical School, London (B.P.R., S.R., R.W.); Department of Psychiatry, Baragwanath Hospital, Soweto, RSA (M.M.-C.); Department of Human Genetics, School of Pathology, South African Institute of Medical Research, University of the Witwatersrand, Johannesburg (T.J.)*

State University of New York - Stony Brook/Oxford/Sequana (SB/Ox/Sq) **Lynn E. DeLisi, Chad Garner, Mary Kelly, Carrie LeDuc, Lon Cardon, Jay Lichter, Tim Harris, Josephine Loftus, Gail Shields, Margarite Comasi, Antonio Vita, Angela Smith, Jay Dann, Geoff Joslyn** *Department of Psychiatry, SUNY Stony Brook (L.E.D., A.S.); Sequana Therapeutics, Inc., San Diego (C.G., M.K., C.L., L.C., J.L., T.H., G.J.); Department of Psychiatry, The Warneford Hospital, Oxford, UK (J.L., J.D.); (4) University of Milan (M.C., A.V.)*

University College London Medical School (UCL) **Hugh Gurling, Gursharan Kalsi, Jon Brynjolfsson, David Curtis, Thordur Sigmundsson, Robert Butler, Tim Read, Patrice Murphy, Andrew Chih-Hui Chen, Hannes Petursson** *Molecular Psychiatry Laboratory, Department of Psychiatry, University College London Medical School (H.G., G.K., T.S., R.B., T.R., P.M., A.C.C.); Department of Psychiatry, Borgarspitalinn, University of Iceland (J.B., H.P.); Joint Academic Department of Psychological Medicine, London Hospital Medical College (D.C.)*

University of Utah **Bill Byerley, Mark Hoff, John Holik and Hilary Coon** *Department of Psychiatry, University of Utah Medical Center, Salt Lake City, Utah.*

US/Australia **Douglas F. Levinson, Derek J. Nancarrow, Raymond R. Crowe, Nancy Andreasen, Jeremy M. Silverman, Richard C. Mohs, Larry J. Siever, Jean Endicott, Lawrence Sharpe, Marilyn K. Walters, David P. Lennon, Nicholas K. Hayward, Lodewijk A. Sandkuijl, Bryan J. Mowry** *Department of Psychiatry, Allegheny University of the Health Sciences, Philadelphia (D.F.L.); Department of Psychiatry, University of Queensland, Wolston Park Hospital, Brisbane (D.J.N., D.P.L., B.J.M.); Queensland Institute of Medical Research, Brisbane (D.J.N., M.K.W., N.K.H.); Mental Health Clinical Research Center and Department of Psychiatry, The University of Iowa College of Medicine (R.R.C., N.A.); Department of Psychiatry, Mt. Sinai School of Medicine, New York (J.M.S., R.C.M., L.J.S.); New York State Psychiatric Institute, Columbia University, New York (J.E., L.S.); Institute of Clinical Genetics, Erasmus University, Rotterdam & Department of Genetics, University of Leiden & Department of Medical Genetics, University of Groningen.(L.A.S.)*

Vienna/Kiel **Harald N. Aschauer, Kurt Meszaros, Elisabeth Lenzinger, Karoline Fuchs, Liu Yang, Angela M. Heiden, Hans W. Moises** *Department of General Psychiatry, University Hospital for Psychiatry, Vienna (H.N.A., K.M., E.L., K.F., A.M.H.); Department of Psychiatry, Kiel University Hospital (L.Y., H.W.M.)*

Additional collaborators **Leonid Kruglyak, Mark J. Daly, Tara C. Matise** *Whitehead Institute for Biomedical Research, Cambridge (USA) (L.K., M.J.D.); Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York (T.C.M.).*

INTRODUCTION

Genetic factors contribute 65-85% of the liability to schizophrenia, probably through polygenic or oligogenic multifactorial inheritance (Gottesman and Shields, 1967; reviewed in Tsuang et al, 1991). Schizophrenia is a relatively common disorder which has a lifetime prevalence of 0.6-1.0% and shows familial co-aggregation with schizo-affective disorder (concurrent psychotic and mood syndromes with persistence of the former), other non-affective psychoses, and schizotypal and paranoid personality disorders (Kendler et al, 1993; Maier et al, 1993; Webb et al, 1993). An intensive worldwide effort has been underway to map schizophrenia-related genetic loci through linkage studies of samples of multiplex pedigrees. While few genome scan projects have been completed to date (Coon et al, 1994; Moises et al, 1995), reports of suggestive evidence for linkage have begun to emerge from the many studies in progress. As simulation studies have shown that very large sample sizes may be required to map genetically complex disorders (Hauser et al, 1996), it has become increasingly clear to workers in this field that single studies are unlikely to provide definitive evidence for linkage, and that collaborative studies may be useful to assess reports of suggestive linkage findings.

In late 1994, two schizophrenia research groups reported to other investigators in the field their suggestive evidence for linkage in three chromosomal regions. Straub et al (1995) subsequently published analyses of sixteen 6p24-22 markers in 265 Irish pedigrees, showing a broad region of positive results by several of their sixteen models with a maximum lod2 of 3.51 at D6S296 (codominant genetic model, broad diagnostic model). Pulver et al (1995) reported results from twelve 3p26-24 markers analyzed in 57 pedigrees by three models with a maximum lod2 score of 2.35 at D3S1283 (dominant affected-only model, narrow diagnostic model) and 58.4% of alleles shared by affected siblings; and sixteen 8p22-21 markers with maximum affected-only lod2 scores of 2.35 (dominant) and 2.20 (recessive) at DS8136 and a peak of 66.7% shared alleles.

The present effort was organized to follow up these suggestive findings. To our knowledge this is the largest multicenter follow-up investigation of linkage findings for a complex disorder yet attempted. Previous multicenter collaborations in Huntington's disease (Huntington's Disease Collaborative Research Group, 1993) and breast cancer (Easton et al, 1993) served as models demonstrating the feasibility and utility of such efforts. There were two previous follow-up studies of a schizophrenia finding on chromosome 22 (Pulver et al, 1994; Gill et al, 1996); and four of the present groups also participated in a two-stage genome scan (Moises et al, 1995). It is expected that other regions of interest will emerge from genome scans being conducted by some of the collaborating groups and others in the field, but the three regions under study have been among the first (in addition to chromosome 22) to show sufficiently positive results to warrant collaborative study.

The present collaboration was formed by initially enlisting all of the groups that participated in the larger of the follow-up studies of chromosome 22 markers (Gill et al, 1996). This ensured that the study would be unbiased with respect to any previous work on chromosomes 3, 6 and 8 (i.e., all groups agreed to participate regardless of whether they had positive, negative or no data for markers in these regions). Two additional groups (St. Mary's and Vienna) joined the collaboration in response to fax and email announcements; neither group had previous results in these regions. Note that several participating groups have published their chromosome 6 data (Moises et al, 1995; Antonarakis et al, 1995; Gurling et al, 1995; Mowry et al, 1995; Schwab et al, 1995; Riley et al, in press; Garner et al, in press) using their own analytic methods, which was permitted by our collaboration agreement, but the analyses reported here have not been published previously.

Key design decisions, reached by consensus, included: (1) Multiple markers were studied in each region, selected from those in the original reports. (2) Primary analyses were limited to a narrow diagnostic model. (3) The New sample (all pedigrees not included in the original positive report for the region) was initially analyzed separately to determine whether there was independent evidence for linkage. (4) It was agreed to conduct a limited number of planned, primary analyses for this initial report to simplify statistical interpretation of the results. These included affected-only single-marker heterogeneity lod score (lod2) analyses under dominant and recessive models (Ott, 1991; Lander, 1988), and a non-parametric multipoint affected sibling pair (ASP) analysis (Kruglyak and Lander, 1995). The MCV group also requested an additional analysis for chromosome 6 using the co-dominant genetic model parameters and broad disease definition that yielded the most positive lod score in their sample (Straub et al, 1995); in addition to MCV, eight groups had made broad (schizophrenia spectrum) diagnoses and agreed to

participate in this analysis, which was expicitly considered secondary. No other analyses of the collaborative dataset have been permitted prior to this initial publication, but the investigators have agreed to permit further subsequent analyses.

MATERIALS AND METHODS

Clinical and molecular methods used by each group have been previously described for each group as follows: Bonn (Schwab et al, 1995), Edinburgh (St. Clair et al, 1989; Blackwood et al, 1996), IOP/Cardiff (Gill et al, 1993), JHU (Pulver et al, 1994), LGN-CNRS (Campion et al, 1994), Iceland/Germany (Moises et al, 1995), MCV (Straub et al, 1995; Kendler et al, 1996), NIMH (Gejman et al, 1993; DeLisi et al, 1987), SB/Oxford/Sequana (DeLisi et al, 1995), St. Mary's (Riley et al, in press), UCL (Kalsii et al, 1995), US/Australia (Mowry et al, 1995), Utah (Coon et al, 1994), Vienna/Kiel (Aschauer et al, 1993). Methods relevant to the collaborative analysis are summarized here.

Subjects

All subjects gave informed consent following regulations for each country and institution. Pedigrees were ascertained by each participating group on the basis of a schizophrenia proband and a constellation of ill relatives (by various diagnostic models) informative for linkage studies, typically 2-5 affecteds in 1-3 generations. Ascertainment schemes were variously designed to identify affected sibships, nuclear families or moderate-sized pedigrees (4-5 affecteds). DNA was obtained from available affected and variable numbers of unaffected relatives, usually including available parents of affecteds and some or all siblings.

Clinical Diagnosis

DSM-IIIR Diagnoses were made by each group using semi-structured research diagnostic interviews and reviews of medical records, except that one group (Riley et al, in press) used consensus DSM-IIIR diagnoses made by clinicians who had demonstrated good diagnostic reliability with a research team. Subjects with DSM-IIIR (American Psychiatric Association, 1987) schizophrenia or schizoaffective disorder (Narrow model) were considered affected for the primary analyses. Five groups limited their diagnostic assessments to Narrow diagnoses. Nine of the groups (those marked by asterisks in Table 1, plus MCV) had also assessed subjects for additional putative schizophrenia spectrum diagnoses, and each of these groups submitted a separate Broad diagnosis file designating subjects as affected if they had DSM-IIIR schizophrenia, schizoaffective or schizophreniform disorder, delusional disorder, psychosis not otherwise specified, or schizotypal or paranoid personality disorder. Narrow diagnoses were used in primary analyses because they were included in all of the studies, demonstrate a higher relative risk ratio and are considered more reliably diagnosed (particularly across groups) than Broad diagnoses. Table 1 describes the clinical sample.

Markers and Distances

The markers used in this study, and the inter-marker distances (shown in centiMorgans in brackets) assumed here, were:

Chromosome 3:
D3S1293 [10] 1283 [0] 1266 [11] 1298.

Chromosome 6:
D6S296 [0.01] 277 [3.99] 470 [9] 259 [7] 285.

Chromosome 8:
D8S261 [4] 258 [4] 133 [6] 136 [15] 283.

All markers were dinucleotide repeat microsatellite markers. They were chosen from the markers studied in the original reports for these regions (Straub et al, 1995; Pulver et al, 1995) to span the most positive linkage results in each region. Distances for chromosome 3 and 6 were taken from the CEPH map (31). For chromosome 6, CEPH data only weakly resolved the order of the first two markers, and the MCV group later reversed the order shown here based on new physical mapping evidence. In the new sample, analysis with GENEHUNTER/NPL (Kruglyak et al, 1996) showed that the most likely number of recombinants between the first-second and second-third markers was similar with either order, so the CEPH order and distances were used (with a small distance assumed between the first two markers because obligate recombinants were observed). (Multipoint ASP results were similar with the two orders.) For chromosome 8, CEPH data did not resolve the order in this region. LINKMAP from the LINKAGE package (Lathrop et al, 1984) was therefore used to define the best order on available CEPH pedigree data, and distances for multipoint analysis were determined by GENEHUNTER/NPL in new sample pedigrees with two typed parents.

Table I: Characteristics of the Collaborative Sample‡

Site	# Aff	# Ped	Aff/ Ped	# Inf	#Pairs	Chromosome 3 A	B	C	D	–Chromosome6– E	F	G	H	I	–Chromosome 8– J	K	L	M	N
Bonn	172	65	2.65	48-53	68-69	+	+	+	+	+	+	+	o	+	+	+	+	o	+
Edinburgh*	50	14	3.57	3-11	12	+	+	+	+	+	o	+	+	+	+	+	+	+	+
Iceland/Germany	22	5	4.40	3-5	7-9	+	o	+	+	+	+	+	+	+	+	+	+	+	+
IOP/Cardiff*	48	13	3.69	10-13	17	o	+	+	+	+	o	+	+	+	+	+	o	+	+
JHU Collab	154	57	2.70	46-52	44	o	o	o	o	+	+	o	+	+	o	o	o	o	o
LGN-CNRS*	123	37	3.32	24-40	50-53	+	+	+	+	+	+	+	+	+	+	+	+	+	+
MCV*	618	265	2.33	156-191	214-224	+	+	+	+	o	o	o	o	o	+	+	+	+	+
NIMH	121	44	2.75	29-41	50-51	+	+	+	+	+	o	+	+	+	+	+	+	o	+
SB/Ox/Sq*	267	100	2.67	51-85	102-103	+	+	+	+	+	+	+	o	+	o	+	+	+	+
St Mary's	63	21	3.00	14-19	13	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UCL*	80	23	3.48	17-21	30-32	+	+	+	+	+	+	+	+	+	+	o	+	+	+
US/Australia*	111	45	2.47	36-39	38-39	+	+	+	+	+	o	+	+	+	+	+	+	+	+
Utah*	40	9	4.44	6-9	14-15	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Vienna/Kiel*	30	15	2.00	9-11	12	+	o	+	+	+	+	+	+	+	+	+	+	+	+
TOT	1899	713	2.66	403-567	677-687														
				Mean heterozygosities:		.82	.75	.74	.87	.85	.79	.80	.79	.70	.77	.71	.79	.81	.81

‡# Aff = # of individuals with DSM-IIIR schizophrenia and schizoaffective disorder.
Ped = # of pedigrees submitted for the collaborative study (includes original and new samples).
Aff/Ped = average number of affecteds per pedigree.
Inf = number of informative pedigrees, defined as lod score <> 0.0000 at theta = 0 -- the range indicates the smallest and largest number of informative pedigrees across chromosome 3, 6 and 8 markers.
Pairs = number of affected sibling pairs (calculated as [n of affected sibs]-1.
Under each chromosome, the letter codes refer to markers: A=D3S1293, B=D3S1283, C=D3S1266, D=D3S1298; E=D6S296, F=D6S277, G=D6S470, H=D6S259, I=D6S285; J=D8S261, K=D8S258, L=D8S133, M=D8S136, N=D8S283. For each marker, "+" indicates that the group submitted genotypes for the marker for the primary extension sample analyses, "o" indicates it did not. Mean marker heterozyosities are shown (average of separate estimates from each group excluding JHU).
Note that the originating groups (JHU and MCV) were excluded from the primary analyses of the chromosome 3+8 and 6 regions respectively, but submitted data for all markers in those regions for the analyses of their own data and of the combined data as shown in Table 2. An asterisk indicates the group contributed data for the schizophrenia spectrum analysis shown in Table 4. References for clinical and laboratory methods for each group are given in the text.

Genotyping

Each group genotyped its available DNA specimens. In some cases a group did not submit genotypes for a particular marker, either because of technical difficulties in analyzing the marker or because of resource limitations. Each lab determined PCR product sizes and made allele designations. Note that all analyses used separate estimations of allele frequencies for each sample, as described below. Thus it was not necessary to determine whether allele designations were consistent across labs. Each group determined allele sizes across all families in their sample (except for JHU, see Linkage Analysis below).

Genetic models for lod score analysis

Two models were selected (q = disease allele frequency; f1, f2, f3 = penetrances for normal genotype, disease heterozygote, and disease homozygote respectively):

Dominant (DOM)
q=0.01, f1=0.000005, f2=0.0005, f3=f2.

Recessive (REC)
q=0.14, f1=0.000005, f2=f1, f3=0.0005

These models were selected by assuming a high penetrance ratio (disease:normal genotype) for affected cases of 0.50:0.005 (100:1), and then dividing penetrances by 1000 to reduce the penetrance ratio for unaffected cases to ≈1.0 (Ott, 1991), so that only affected cases contribute significantly to the lod score ("affected-only" analysis). This was preferred for a complex disease because penetrance for a given locus may be too low for an "unaffected" diagnosis to be reliable. Linkage analysts for MCV and JHU considered these models acceptable tests of their findings under a Narrow diagnostic model.

The secondary Broad model analysis (chromosome 6) used a modificiation of the PEN model which yielded the highest lod score in the original MCV analysis (Straub et al, 1995): q = 0.032, f1 = 0.0064, f2 = 0.375 and f3 = 0.75 (with female penetranes set at 50% of these values), modified here by dividing all penetrances by 1000 to create an affected-only model.

Linkage analysis

LINKAGE-format files of pedigree, diagnostic and genotyping data for each dataset (by chromosome) were analyzed with PEDMANAGER (M.P. Reeve, unpublished) for mendelization errors and also for estimation of allele frequencies separately for each sample (the program counts alleles among pedigree founders, after reconstructing missing genotypes for founders when this can be done with certainty). Two-point lod scores (disease vs. marker) were computed with the FASTLINK (Cottingham et al, 1993) version of MLINK (Lathrop et al, 1984), for DOM and REC models for each sample separately (using the allele frequencies calculated for that sample), at recombination frequencies (θ) between 0 and 0.48 at intervals of 0.02. Lod scores for each pedigree were then assembled into files (by marker) for the New, Original and Combined samples for computation of heterogeneity lod scores (lod2; admixture test) by HOMOG (Ott, 1991) (Table 2). Thus, while the overall lod2 score was computed from lod scores for pedigrees from all samples, those pedigree lod scores were calculated using the allele frequencies calculated separately for the sample from which the pedigree was drawn.

Single-marker ASP analyses (Table 3) were carried out with SPLINK (Holmans and Clayton, 1995) which calculates a maximum lod score (MLS) from an estimate of the most likely proportions of sibling pairs with 0, 1 or 2 alleles shared identical-by-descent, using the "possible triangle" method which increases the power of ASP analysis (Holmans, 1995). The SPLINK analyses utilized the separate allele frequency estimates for each sample in calculating sharing proportions for each sample, and then estimated maximum sharing for the entire New and Combined samples. The SPLINK analysis was planned as a check on the consistency of multipoint results, with the latter considered the primary analyses. However, it was decided not to perform multipoint analyses on chromosome 3 data after observing lod2 scores of < 0.10 and single-marker MLS results of 0-0.54 at all markers in the New sample (lod2 and MLS scores cannot be < 0), such that a substantially higher multipoint result would not be expected.

Multipoint ASP analysis was carried out for chromosomes 6 and 8 with MAPMAKER/SIBS (Kruglyak and Lander, 1995). Like SPLINK, SIBS estimates the maximum likelihood proportions of 0, 1 and 2 alleles shared by pairs of sibs across the dataset, making no assumption about dominance variance, and with no genetic model, using the "possible triangle" method, but incorporates rapid estimation of the multipoint distribution to permit calculation of the MLS at every point across a specified map of markers. The SIBS analyses (like SPLINK) weighted the contribution of sibships by a factor of 2/N (Suarez and Hodge, 1979), so that each sibship contributed the equivalent of s-1 pairs (s=number of affected sibs) as an approximate correction for non-independent pairs.

SIBS analyses utilized the separate allele frequency estimations for each sample as follows: datafiles were created for each chromosome containing the entire New or Combined sample, with allele designations for each sample recoded into non-overlapping bins (e.g., alleles 1-20 for Edinburgh, 21-40 for CNRS, etc.) in the pedigree and locus parameter files (totalling 280-302 alleles per marker). The locus parameter file listed the allele frequencies calculated separately for each sample for the appropriate allele designations for that sample, with allele frequencies summing to the number of groups included the analysis.

For chromosome 6 analyses using a Broad diagnostic model, nine groups (MCV plus the other eight groups marked by asterisks in Table 1) submitted files with these diagnoses designated as affected. The MCV group computed lod2 scores (using MENDEL [Lange et al, 1988]) and C-test (single point) and D-test (multipoint) tests for linkage with heterogeneity (MacLean et al, 1992), under the DOM and REC models (Table 4) and the affected-only version of the PEN model. Lod2 scores were also computed under DOM and REC for Narrow diagnoses for these groups separately (with MLINK and HOMOG) for comparison (Table 4).

Results are reported as lod2 and SIBS or SPLINK maximum lod scores (MLS), with nominal (asymptotic, point-wise) p values. The reader should note that lod2 (heterogeneity lod scores) are two-parameter tests (recombination fraction and proportion of linked families) which are approximately equivalent to homogeneity lod scores after deducting a correction of about 0.3 (Ott, 1991), but which actually follow an unusual distribution described by Faraway (1993) who also provided the appropriate method for calculating nominal p values (Faraway, personal communication): the lod2 score is converted to a chi-square value with 1 df by multiplying by 4.6 (2 x logn[10]); the area to

Table II: Lod2 and SIBS Maximum Lod Scores: Chromosomes 6 and 8‡

		——DOM——				——REC——				——SIBS——	
	Peds	lod2	p	α	θ	lod2	p	α	θ	Pairs	MLS/ (shared)
				Chromosome 6							
New										463	*2.19*
D6S296	354	*0.00*		1.00	>0.5	*0.47*		0.08	0.00		(.559)
D6S277	261	*0.51*		0.11	0.00	*1.22*	*.018*	0.15	0.00		*p=.001*
D6S470	331	*0.58*		0.18	0.12	*1.20*		0.12	0.00		
D6S259	241	*0.05*		0.04	0.00	*0.58*		0.12	0.00		
D6S285	357	*0.31*		0.09	0.02	*1.19*		0.59	0.24		
MCV										224	*1.03*
D6S296	197	*1.31*		1.0	0.26	*1.31*		0.25	0.04		(.547)
D6S277	181	*1.14*		1.0	0.24	*1.20*		0.22	0.00		*p=.02*
D6S470	212	*1.38*	*.012*	0.93	0.24	*0.61*		1.00	0.32		
D6S259	220	*0.48*		0.37	0.18	*0.11*		1.00	0.38		
D6S285	210	*0.39*		1.0	0.32	*0.30*		0.31	0.20		
Combined										687	*2.68*
D6S296	551	*0.06*		0.03	0.02	*1.47*		0.12	0.00		(.547)
D6S277	442	*1.20*		0.15	0.00	*2.35*	*.001*	0.18	0.00		*p=.0004*
D6S470	543	*1.66*		0.26	0.12	*1.66*		0.20	0.10		
D6S259	461	*0.34*		0.08	0.02	*0.61*		0.09	0.00		
D6S285	567	*0.56*		0.11	0.04	*1.45*		0.53	0.24		
				Chromosome 8							
New										633	*1.58*
D8S261	384	*0.73*		0.96	0.34	*2.22*	*.0014*	0.17	0.00		(.541)
D8S258	418	*0.34*		0.08	0.00	*0.78*		0.23	0.14		*p=.005*
D8S133	381	*0.97*		1.00	0.32	*1.63*		0.20	0.06		
D8S136	359	*0.99*		0.74	0.30	*1.23*		1.00	0.32		
D8S283	463	*0.50*		1.00	0.40	*0.60*		1.00	0.36		
JHU										44	*2.90*
D8S261	45	*0.61*		0.31	0.00	*1.28*		0.91	0.16		(.704)
D8S258	38	*1.10*		0.79	0.16	*1.62*		0.70	0.08		*p=.0002*
D8S133	42	*1.35*		0.44	0.00	*1.84*		0.85	0.10		
D8S136	44	*2.24*	*.0013*	1.0	0.14	*2.11*		1.0	0.12		
D8S283	41	*0.30*		1.0	0.30	*0.91*		0.27	0.00		
Combined										677	*2.73*
D8S261	429	*0.98*		0.38	0.22	*3.06*	*.00018*	0.19	0.00		(.553)
D8S258	456	*0.85*		0.23	0.14	*1.62*		0.27	0.12		*p=.0003*
D8S133	423	*1.60*		0.81	0.28	*2.73*		0.25	0.06		
D8S136	403	*2.08*		1.0	0.30	*2.19*		0.90	0.28		
D8S283	504	*0.67*		1.0	0.38	*1.03*		0.13	0.06		

‡ Shown for dominant and recessive models are lod2 scores for each marker and corresponding values of α and θ. Peds=the number of informative pedigrees (lod<>0.0000) for each marker. Shown for MAPMAKER/SIBS analyses are the number of typed pairs (where sibships are weighted to contribute s-1 pairs), and the maximum lod score (MLS) and maximum estimate of proportion of shared alleles (see Figure 1). The maximum lod2 and MLS result for each dataset is bolded, and the nominal p-value shown.

the left of this value is then squared and subtracted from 1; and this number is divided by 2 to obtain the nominal p value. For MLS scores, approximately equivalent asymptotic p values were

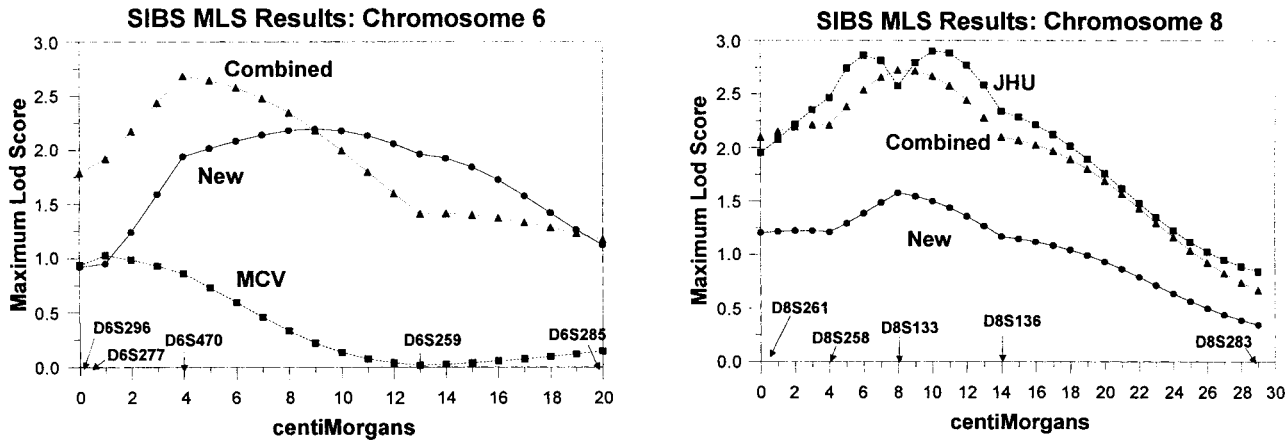

Fig. 1: SIBS MLS results for chromosomes 6 and 8. Shown are multipoint MLS values (MAPMAKER/SIBS) at intervals of 1 cM across the map of markers for each chromosome. Results for the New (collaborative) sample are shown in circles, the Original sample (JHU or MCV) in squares, and for Combined (New+ Original) sample in triangles. See Table II for N and peak values.

Table III: Results of SPLINK Single-Marker ASP Analyses for the New Sample‡

	A	B	A	B	A	B	A	B	A	B
					Chromosome 3 - New Sample					
Marker	D3S1293		D3S1283		D3S1266		D3S1298			
N	388.9	171.4	354.5	163.0	366.0	162.7	409.0	178.6		
MLS	*0.07*	*0.23*	*0.02*	*0.00*	*0.00*	*0.00*	*0.54*	*0.24*		
p(0,1,2)	.25,.49,.26	.24,.48,.28	.25,.50,.25	.25,.50,.25	.25,.50,.25	.25,.50,.25	.24,.48,.28	.24,.48,.25		
					Chromosome 6 - New Sample					
Marker	D6S296		D6S277		D6S470		D6S259		D6S285	
N	331.9	202.4	224.2	135.0	286.2	173.7	202.7	104.5	278.7	170.1
MLS	*0.79**	*1.54***	*1.28***	*1.31***	*1.42***	*1.60***	*0.34*	*0.58*	*0.96**	*1.20**
p(0,1,2)	.23,.47,.30	.22,.44,.33	.20,.48,.32	.22,.44,.34	.20,.49,.31	.22,.44,.34	.24,.47,.29	.20,.48,.32	.22,.48,.30	.19,.49,.32
					Combined Sample					
N	438.0	217.8	356.1	163.3	416.6	194.4	337.5	128.4	401.4	192.0
MLS	*1.19**	*1.58***	*2.15***	*1.17**	*1.55**	*1.24**	*0.28*	*1.00**	*1.08**	*1.54**
p(0,1,2)	.23,.47,.30	.22,.44,.33	.22,.45,.32	.22,.44,.33	.22,.47,.31	.22,.45,.33	.24,.48,.28	.19,.49,.32	.22,.48,.30	.20,.47,.33
					Chromosome 8 - New Sample					
Marker	D8S261		D8S258		D8S133		D8S136		D8S283	
N	352.7	154.8	313.4	141.4	348.4	155.4	322.6	111.5	424.0	196.9
MLS	*0.97**	*0.44*	*0.73*	*0.17*	*1.74****	*0.88**	*0.99**	*0.67*	*0.36*	*0.07*
p(0,1,2)	.23,.47,.30	.22,.49,.29	.23,.47,.30	.24,.48,.28	.23,.45,.32	.22,.46,.32	.21,.50,.29	.23,.45,.32	.24,.48,.28	.24,.49,.27
					Combined Sample					
N	384.3		342.8		376.6		354.0		460.0	
MLS	*1.69****	*1.29**	*1.46***	*0.84**	*2.71****	*2.32****	*1.63***	*1.41***	*0.66*	*0.32*
p(0,1,2)	.22,.46,.32	.20,.47,.33	.21,.48,.31	.22,.46,.32	.22,.44,.33	.19,.45,.36	.21,.48,.31	.21,.43,.35	.24,.48,.29	.24,.48,.29

‡For each marker, column A gives results for all pairs, and column B for pairs with both parents typed. The latter is a check on the dependency of the MLS on estimates of allele frequencies, which would be suggested if sharing proportions were lower when parents were typed; here, proportions are generally similar. N = Number of fully-informative sib pairs that would yield the amount of linkage information observed. MLS = maximum lod score, computed with SPLINK separately for each marker, based on maximum likelihood sharing proportions, using the possible triangle method. p(0,1,2) = the maximum likelihood proportions of pairs with 0, 1 or 2 alleles shared IBD. * $p < .05$, ** $p < .01$, *** $p < .005$

Table IV: Maximum Two-point Lod Scores (Recessive) by Group (New Sample)

	----------Chromosome 6----------					----------Chromosome 8----------				
Marker:	296	277	470	259	285	261	258	133	136	283
Bonn	0.15	-0.01	0.79		***1.48***	0.08	**0.33**	0.00		0.03
Edinburgh	0.01		0.01	***0.42***	0.00	0.42	0.65	0.03	***1.10***	0.01
Iceland/Germany	0.14	-0.01	-0.01	-0.01	-0.01	***0.57***	0.12	-0.01	-0.01	0.43
IOP/Cardiff	-0.01		***0.57***	0.22	0.00	0.11	0.00		0.00	0.00
JHU Collab	0.07	***0.86***		0.12	0.07					
LGN-CNRS	0.19	0.60	0.00	***0.92***	0.03	0.31	0.07	0.15	0.09	***0.40***
MCV						0.23	0.43	1.04	***1.32***	0.06
NIMH	-0.01		0.04	0.00	0.13	0.00	***0.94***	0.12		0.00
SB/Ox/Sq	0.19	0.30	0.16		***0.64***		-0.01	0.04	0.00	0.00
St Mary's	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.11
UCL	0.00	0.05	***0.31***	-0.01	0.10	0.21		***1.04***	0.49	0.73
US/Australia	0.14		0.19	***0.50***	0.26	***0.54***	0.00	0.04	0.00	0.00
Utah	0.00	0.00	0.00	0.00	0.00	0.00	***0.75***	0.01	0.18	0.00
Vienna/Kiel	***1.13***	0.13	0.14	0.09	0.01	0.12	0.00	0.10	0.00	***0.63***
Lod2(max)	0.47	***1.22***	1.20	0.58	1.19	***2.22***	0.78	1.63	1.23	0.60

Shown are the maximum two-point lod scores (recessive model, narrow diagnoses, assuming homogeneity) for each New sample group separately, by marker, to illustrate the relative contribution of each sample to positive scores. Maximum lods occur at different values of theta (not shown) in the different groups. The highest lod(max) for each group (from Table 2) on each chromosome is bolded, if it was > 0.3. Also shown are the maximum lod2 scores for each marker across all pedigrees (New sample) -- note that the lod2 is not a sum of the lod(max) values, but is shown to indicate the total evidence for linkage at each marker.

calculated as described by Holmans (1993). We considered carrying out formal power analyses through simulation studies, but concluded that it would be prohibitive to carry out appropriate simulations such a large sample of various pedigree structures, with highly polymorphic markers, for the primary methods ofanalysis (lod2 and multipoint SIBS). Nominal thresholds of significance should be reasonably accurate due to the large sample sizes.

Note that the JHU group had taken a different approach to allele frequencies, as follows: alleles had been sized within families only. If x alleles were observed within a single family, they had been designated with integers from 1-x. If the largest number of alleles observed in any one family was y, then all alleles for that marker were assumed to have a frequency of 1/y (equal frequencies). To accomodate this approach, these frequencies were also used for the primary analyses (SIBS and lod2 analyses by the narrow diagnostic model; JHU did not submit Broad diagnoses). For SPLINK, the program used for the multi-group analysis calculated allele frequencies automatically for each subsample, such that the proportion of "1" alleles would have been computed regardless of the fact that "1" referred to different alleles in different families. As a check on this, the MLS for the JHU group was calculated separately with these calculated frequencies and with equal frequencies, and differences between the two analyses were small. Thus the MLS results reported here reflect the (incorrectly) computed allele frequencies for the JHU sample. Again, for all other samples, allele frequencies were computed from founders (as described above) on the basis of cross-family allele designations, and the separate allele frequencies for each sample were used in all linkage tests.

RESULTS

Primary Analyses: Additional Support for Linkage on Chromosomes 6 and 8

The characteristics of the sample are shown in Table I. Lod2 and SIBS MLS results are shown in Table II, and SIBS results for chromosome 6 and 8 are shown graphically in Figure 1. Single-marker SPLINK results are shown in Table III for comparison with the multipoint results, but it should be appreciated that the multipoint analysis provides more information and was planned as the primary non-parametric analysis of these data. Because all chromosome 3 markers produced lod2 scores $\leq$ 0.08 and single-marker SPLINK MLS results between 0 and 0.54, SIBS multipoint

Table V: Lod2 and C-test Analyses using Broad Diagnostic Model (Chromosome 6)*

	Narrow-lod2		Broad-DOM		Broad-REC		Broad-PEN	
	DOM	REC	lod2	*C-dev*	lod2	*C-dev*	lod2	*C-dev*
New8								
D6S296	0.01	0.66	0.09	*0.11*	0.98	*1.83*	0.38	*0.52*
D6S277	0.94	1.17	0.64	*1.43*	0.41	*0.91*	0.38	*0.94*
D6S470	0.17	0.74	0.02	*-.38*	0.01	*-.06*	0.01	*-0.60*
D6S259	0.15	1.47	0.15	*0.94*	**1.36**	***2.22***	0.34	*0.96*
D6S285	0.11	0.54	0.11	*0.53*	0.60	*1.68*	0.17	*0.78*
MCV								
D6S296	1.31	1.31	**2.83**	*3.10*	2.51	***3.70***	2.70	*3.40*
D6S277	1.14	1.20	1.55	*2.10*	0.64	*1.90*	1.39	*1.70*
D6S470	1.38	0.61	2.26	*3.10*	1.03	*1.50*	1.72	*2.20*
D6S259	0.48	0.11	1.01	*1.70*	0.61	*0.60*	0.78	*0.60*
D6S285	0.39	0.30	1.25	*1.70*	1.77	*2.70*	1.45	*1.80*
New8+MCV								
D6S296	1.78	3.27	1.79	*2.20*	2.89	***3.90***	**2.90**	*2.80*
D6S277	2.16	2.45	1.38	*2.40*	.97	*1.90*	1.48	*2.00*
D6S470	1.92	2.57	1.15	*1.60*	.55	*0.97*	1.14	*1.30*
D6S259	0.91	1.24	.94	*1.90*	1.76	*1.70*	1.02	*1.50*
D6S285	0.72	1.27	.78	*1.60*	2.18	*3.10*	1.26	*1.90*

*Shown are results of secondary analyses of chromosome 6 data under a Broad diagnostic model, and DOM, REC and PEN transmission models, for New8 (8 groups, marked by asterisks in Table 1, which submitted Broad diagnoses), MCV alone, and New8+MCV combined. The Narrow lod2 results are shown for comparison purposes only, to permit the reader to examine the effect of adding Broad model cases to the analysis of the New8 subsample -- Table 2 shows the Narrow model results for the entire sample (planned primary analyses). The New8 subsample included 129-219 informative pedigrees by Narrow and 160-260 by Broad diagnoses, with markers D6S277 and D6S259 having reduced numbers as shown in Table 1.

For each marker, lod2 (a likelihood test) and C-test deviate (a simulation-based test of linkage with heterogeneity) have been computed. Nominal p-values (calculated as described in Methods for lod2 scores, and computed by simulation for C-deviates) for the maximum result for each sample are:

New8 - lod2 (REC) = 1.36 (p=.012), C-dev (REC) = 2.22 (p=.013);
MCV - lod2 (DOM) = 2.83 (p=.00031), C-dev (REC) = 3.70 (p=.0001);
New8+MCV - lod2 (PEN) = 2.90 (p=.0026), C-dev (REC) = 3.90 (p=.0001).

analysis was not carried out for this chromosomal region.

For chromosome 6, the most positive results were a SIBS MLS of 2.19 (nominal p = .001) in the New sample and 2.68 (nominal p = .0004) in the Combined sample. (For comparison, the standard homogeneity lod scores with equivalent p-values would be 2.00 and 2.40). Note that the MCV sample's lod2(max) of 1.38 is consistent with their published results using a narrow diagnostic model, and much lower than their maximum scores using a Broad model (Straub et al, 1995) as discussed below.

For chromosome 8, the most positive results were lod2 of 2.22 by a recessive model at D8S261 (p = .0014) in the New sample and 3.06 by a recessive model at D8S261 (p = .00018) in the Combined sample. (Similar p-values would be associated with homogeneity lod scores of 1.94 and 2.78) . The JHU dataset produced a lod2 score of 2.24 by a dominant model at D8S136 (p = .0013, similar to that of a homogeneity lod score of 1.96).

How consistent were the findings across groups? The contribution of each group's sample to the maximum lod scores cannot be determined exactly because the relevant statistics are maximized over the entire dataset. As an illustration of the variation among samples, Table IV shows the maximum two-point homogeneity lod score for each sample at each marker (by the recessive model, which yielded the highest lod scores in the new sample for both chromosomes). For reference, the maximum lod2 result for each marker is also shown. Slightly positive lods were observed in most of the subsamples (> 0.3 in 10 of 13 samples for chromosome 8 and 9 of 13 samples for chromosome 6), but there was little consistency in the location of the maximum lod score.

A broad diagnostic model did not increase evidence for linkage

Table V shows the results of the planned secondary analysis of chromosome 6 data using a Broad diagnostic model (with data under the Narrow model shown for comparison) with three genetic models (DOM, REC, PEN) and two different test statistics, lod2 and the C-test deviate (MacLean et al, 1992). Evidence for linkage for the MCV sample is lower here than in their published report (Straub et al, 1995), with the difference attributable to the present study's use of affected-only models. The MCV group's hypotheses (based on their data) were that higher lod2 scores would be observed under the Broad (vs. the Narrow) diagnostic model, the PEN (vs. DOM or REC) transmission model, and with the simulation-based C-test (vs. the likelihood-based lod2 test). The pedigrees from the 8 groups contributing to this analysis, other than MCV, are here designated as New8 sample. In the New8 sample, lod2 scores for dominant and recessive models and a Broad diagnostic model were always lower than lod2 scores for the same genetic model and a Narrow diagnostic model. As expected from published results, the highest lod2 and C-deviate scores for the MCV sample were observed with the PEN model and a Broad diagnostic model. The highest lod2 and C-deviate scores in the combined (New8 + MCV) sample were for the PEN and REC models respectively under a Broad diagnostic model. For the PEN model, this was almost entirely due to the MCV sample's contribution. For REC, the combined lod2 score of 2.89 at D6S296 is lower than the lod2 score of 3.27 observed in these 9 samples for REC under a Narrow diagnostic model (but note that only primary analysis results for the entire sample shown in Table II should be considered in interpreting the significance of the findings for the Narrow diagnostic model).

DISCUSSION

The result of the primary analyses for chromosomes 6 and 8 can be interpreted in many ways. It is expected that some of the collaborators will carry out additional analyses and publish their own results and interpretations. Discussions among the present collaborators have demonstrated a consensus on the following two points:

1. The results support the need for additional efforts to identify or exclude schizophrenia susceptiblity genes on chromosomes 6 and 8, where linkage analyses using conservative assumptions (affected-only analyses of narrowly-diagnosed cases) in a large, new, unbiased sample yielded positive evidence of a similar order of magnitude as the original samples. This result stands in sharp contrast to the negative results observed in most previous attempts to replicate schizophrenia linkage findings (Baron, 1995), and in our own chromosome 3 data.

2. Although there is no consensus about the statistical interpretation of linkage results for complex disorders, it is clear that the present results fall short of any standard of "significant linkage." For example, the traditional threshold of significance for mendelian disorders has been a homogeneity lod score of 3.0 (nominal p = .0001), based on a set of assumptions about prior and posterior probabilities in linkage tests for mendelian disorders (Ott, 1991). None of the test statistics for the planned, primary analyses in the present study reached this level of significance. There is no consensus about defining thresholds of statistical significance for likelihood ratio tests in the many situations in which they are currently applied.

Discussion of these results among the present collaborators has not produced agreement on a more precise statistical interpretation. One approach would be to follow the proposal of Lander and Kruglyak (1995) that all linkage tests for complex disorders be interpreted in relation to the likelihood of observing a result by chance in a genome-wide search. They presented calculations showing that nominal point-wise p values of .000049 (for lod score analysis) or .000022 (for sib-pair analysis) would be observed by chance in 5% of complete genome scans using multipoint analyses, and they suggest adopting these values as approximate thresholds for "significant" linkage. Similarly, p values of .0017 and .00074 would be expected by chance approximately once per genome scan, which they suggest as a threshold for "suggestive" linkage. By this model, p values of .01 would be considered confirmation of previous significant (but not suggestive) linkage results. These criteria for suggestive linkage are exceeded by the maximum results for chromosome 6 (Combined sample) and chromosome 8 (New and Combined samples). However, the initial JHU result for chromosome 8 was at the level of suggestive linkage; and the initial MCV result for chromosome 6, while very close to the proposed threshold for significant linkage (lod2 = 3.51, nominal p=.000059), would require some degree of adjustment because multiple analytic models were tested in that study. Thus by the proposed guidelines, our finding of "suggestive" linkage in both the New and Combined samples, in the absence of a clearly "significant" prior result,

cannot be considered confirmation or replication of linkage. Furthermore, a number of issues concerning the applicability of this model are still under discussion (Witte et al, 1996; Curtis, 1996; Lander and Kruglyak, 1996) including whether such discrete thresholds are most useful, how to correct for the use of several partially-correlated tests, how to apply the criteria in following up previous findings of varying magnitudes, and whether they should be applied to two-point analyses (such as our lod2 tests) as opposed to multipoint analyses. (It should be noted that the DOM and REC lod2 and SIBS analyses of the the New sample under the Narrow diagnostic model were planned as primary analyses, and that these three tests are clearly non-independent in these data.)

The absence of definitive proof or disconfirmation of linkage should not be surprising, despite the large sample size. Simulation studies (Suarez et al, 1995) suggest that in a polygenic disease, the initial detection of linkage may occur because one of the susceptibility loci happens (for stochastic reasons) to appear more strongly linked with disease in one sample than in the general population. Replication is then difficult because (again due to stochastic variation, and not because of any difference in genetic effect) linkage results are likely to be weaker in other samples. Thus the present results could represent chance findings, or they could reflect the kind of weak linkage one might expect in a replication study of a complex disorder. Several other methodological problems could also be reducing the observed strength of linkage. Clinical diagnosis (even with stringent research methods) probably remains an approximation of the actual phenotype, given the reliance on psychotic subjects' self-report and on partial information from clinical records, the lack of a gold standard to validate clinical diagnoses (and the virtual impossibility of achieving diagnostic uniformity across a large number of studies, countries and languages, given the labor-intensive nature of procedures for making and reviewing diagnoses), the absence of known familial subtypes to subdivide cases for analysis, and the dearth of knowledge about the features and prevalence of non-genetic cases. The most tightly-linked markers might not have been studied, given the limited number of markers to which each group could commit to completing and the decision to select markers from the original reports. Undetected genotyping errors are more likely in the small (and often incomplete) pedigrees seen in studies of complex disorders, and in analysis of dinucleotide repeat markers. Finally, the small number of quite conservative analyses might have underestimated the degree of linkage -- although, given the marginal magnitude of the findings, use of multiple tests would have raised additional questions about the significance of the results.

Sample-by-sample variation in evidence for linkage can be observed with weak linkage or false positive results, as well in the face of true genetic heterogeneity among samples (although there was no statistical evidence for this here) or because of technical problems as discussed above. With regard to possible gene location, both chromosomes demonstrated considerable variation by sample in the location of maximum positive scores (the reason for which is not known). For chromosome 8 the location of the peak multipoint SIBS results was similar for the entire new sample and the JHU sample.

The present data fail to provide independent confirmation of the MCV group's finding of stronger evidence for linkage using a Broad diagnostic model, or using a codominant (PEN) model. There are a number of possible explanations for this: the MCV group may be more accurately diagnosing genetically-related spectrum disorders; the Irish sample could differ from the others in the spectrum of phenotypes; the difference between the Narrow and Broad models in the MCV sample could have been due to chance; or the chromosome 6p finding itself could be a false positive.

A more general issue is whether multicenter follow-up studies of positive findings can play an important role in the identification of susceptibility genes for complex disorders. Collaborative analyses played a substantial role in the cloning of genes for major locus disorders including Huntington's disease (Huntington's Disease Collaborative Research Group, 1993) and familial breast cancer (Easton et al, 1993). In complex disorders, relatively large collaborations have been organized for two-stage genome searches, using a low threshold of significance in the first stage to select regions for a second (collaborative) stage, but typically a minority of samples in the world were included; examples are Davies et al (1994) in insulin-dependent diabetes mellitus (IDDM) (282 sib pairs), and Moises et al (1995) in schizophrenia (65 pedigrees, including some large pedigrees). Two collaborations were organized to follow up the JHU group's finding (Pulver et al, 1994a) of a lod2 of 1.54 on chromosome 22 in a genome screen: JHU and three other groups (256 families) typed three markers around IL2RB, with negative results (Pulver et al, 1994b); and in a collaboration organized by the Institute of Psychiatry group, JHU and ten other groups typed a single additional

marker (D22S278, chosen because several groups had positive findings nearby) in 436 sibships, with positive but equivocal results (MLS = 1.79) (Gill et al, 1996). The JHU group recently reported several schizophrenia patients with microdeletions within the nearby velo-cardio-facial syndrome (VCFS) locus (Karayiorgou et al, 1995), and schizophrenia-like syndromes in 4 of 14 adult VCFS patients examined (Pulver et al, 1994c), suggesting a possible etiological role for this locus in a small proportion of schizophrenia cases. For bipolar disorder, five groups have contributed data to the 1996 Genetic Analysis Workshop meeting for study of a chromosome 18 bipolar disorder finding (Berrettini et al, 1994).

Well-designed and timely multicenter follow-up studies can help to direct research efforts toward promising findings (like chromosomes 6 and 8 in schizophrenia) and away from those (like chromosome 3) which disappear in a larger sample. They can also lend credibility to findings and to the field itself, by focussing attention on the difficult and slow task of gradually honing in on susceptibility genes, a task that has been hampered by repeated waves of excessive publicity for suggestive findings in individual samples followed by disillusionment (even among fellow scientists) when rapid replication fails to materialize. Whether multicenter studies will in fact play a critical role remains to be seen. For example, several groups have begun efforts to localize and clone schizophrenia susceptibility genes on chromosomes 6, 8 and 22, on the basis of available suggestive evidence for linkage. Thus for complex disorders like schizophrenia, linkage studies will sometimes be used only to obtain initial clues about gene location, and investigators may turn to other methods (linkage disequilibrium, screening for candidate genes, mutational analysis) before obtaining the kind of definitive evidence for gene location that would be expected from linkage studies alone in mendelian disorders. Multicenter analyses can still play a role in this process: if initial cloning efforts are successful, then multicenter analyses can assist in detection of additional loci and clarification of genetic mechanisms and clinical correlations; if they are not, then multicenter analyses can help to identify the subsequent findings most worthy of further study. We would also note that, given the limitations of multicenter studies such as differences in diagnostic and laboratory methods, the collection of large single samples should continue to be a priority for diseases such as schizophrenia.

We suggest that timely multicenter follow-up studies of linkage findings can be useful to the field if adequate attention is paid to aspects of study design such as formation of a large and unbiased sample; balance between central coordination and collaborative decision-making to maintain trust in the process; study of an adequate number of the best available markers spanning a region of interest; and thoughtful selection of a genetic analysis strategy. Continued improvements in diagnostic methods, epidemiological knowledge and genotyping accuracy may be needed to strengthen findings in the face of complex, poorly understood genetic mechanisms.

In conclusion, the consensus of the present investigators is that the present results provide additional suggestive evidence, but not statistically significant evidence, to support the hypotheses of the presence of schizophrenia susceptibility genes in these regions of chromosomes 6 and 8 (but not 3). Several of the present investigators have viewed the results with more skepticism because study of a considerably enlarged sample provided no greater support than the original reports, suggesting either false positive results or weak gene effects. We currently have limited methods to distinguish between these possibilities, given the data cited above from simulation studies that stochastic variation in linkage results for polygenic disorders may lead to inconclusive results even in samples larger than the present one, and the many methodological factors that could weaken linkage results in this type of study. One point favoring a more positive interpretation is that two of three regions studied on the basis of prior hypotheses yielded suggestively positive results, which would be much less likely to occur by chance than zero or one such result. Thus we conclude that our evidence for linkage on chromosomes 6 and/or 8 is too strong to ignore, but too modest to resolve the status of these regions. Perhaps the most revealing result of the present study is that most of the investigators report it has increased rather than decreased their interest in efforts to identify susceptibility genes in these regions by linkage and other methods.

ACKNOWLEDGEMENTS

This study was coordinated by D. F. Levinson. The analyses were carried out by D. F. Levinson, P. Holmans, L. Kruglyak, M. J. Daly, S. Easter, C. MacLean and T. C. Matisse. The study coordinator would like to express appreciation to the contact persons for each group, to those who assisted with analyses, to those who joined in the lively and useful discussions that shaped this manuscript, and to Dr. Ann Pulver for suggesting this multicenter

study and for her advocacy of collaborative follow-up studies in schizophrenia research. Support for this study was provided in part by NIMH grants MH41953, MH45390 and MH01277 (MCV); MH45097 and KO2-01207 (US/Australia - Philadelphia); MH43271, MH31593, MH00625, and MH00735 (US/Australia - Iowa); MH44212, MH01089, MH10168 (Utah); MH42827, MH48858, MH45212, and DK31775 (US/Australia - Mt. Sinai); and MH44292 (J. Ott); NIAA grant AA08402 (US/Australia - Iowa); VA grant VA-4125-019 (Mt. Sinai); the (UK) Medical Research Council (Cardiff; Edinburgh; UCL [grant G880473N]); the Scottish Rite Schizophrenia Research Program (Utah); German Research Foundation grants M0429 (Iceland/Germany); AL 230/2-2, MA 1142/4-3 and WI 453/7-4 (Bonn); Wolfson Foundation Grant AJH2108 (St. Mary's); the National Health and Medical Research Council of Australia, Rebecca L. Cooper Medical Research Foundation, Queensland Department of Health, and Australian NHMRC Network for Brain Research into Mental Disorders (US/Australia); Austrian Research Foundation grant 7639 (Vienna); the Wellcome Trust (Edinburgh; UCL); the Icelandic Science Council, Schizophrenia and National Emergency, the Joseph Levy Charitable Foundation, and the Priory Hospital (UCL); and the German-Israeli Foundation for Scientific Research (Bonn). Assistance from Genethon is gratefully acknowledged (LGN-CNRS; Iceland-Germany).

AVAILABILITY OF DATA

The investigators will make raw data from this study available for further analysis under certain conditions (including demonstration of qualifications in genetic analysis, advance discussion of the planned use of the data, and assurance of confidentiality of pedigree structures). Unique confidentiality and other issues in the study of psychiatric disorders necessitate consideration of such requests on a case-by-case basis, but every effort will be made to cooperate with qualified investigators. Requests should be made through one of the investigators or the study coordinator (levinson@allegheny.edu).

REFERENCES

Antonarakis SE, Blouin JL, Pulver AE, et al. (1995): Schizophrenia susceptibility and chromosome 6p24-22. Nature Genetics 11:235-6.

Aschauer HN, Fischer G, Isenberg KE, et al. (1993): No proof of linkage between schizophrenia-related disorders including schizophrenia and chromosome 2q21 region. Eur Arch Psychiatry Clin Neuroscience 243:193-8.

Baron, M (1995): Genes and psychosis: old wine in new bottles?. Acta Psychiatrica Scandinavica. 92:81-6.

Berrettini WH, Ferraro TN, Goldin LR, et al. (1994): Chromosome 18 DNA markers and manic-depressive illness: evidence for a susceptibility gene. Proc Nat Acad Sci USA 91:5918-21.

Blackwood DH, He L, Morris SW, et al. (1996): A locus for bipolar affective disorder on chromosome 4p. Nature Genetics 12:427-30.

Campion D, d'Amato T, Bastard C, et al. (1994): Genetic study of dopamine D1, D2, and D4 receptors in schizophrenia. Psychiatry Research 51:215-30.

Coon H, Jensen S, Holik J, et al. (1994): Genomic scan for genes predisposing to schizophrenia. American Journal of Medical Genetics. 54:59-71.

Cottingham RW Jr, Idury RM, Schaffer AA (1993): Faster sequential genetic linkage computations. American Journal of Human Genetics 53: 252-263.

Curtis D (1996): Genetic dissection of complex traits. Nature Genetics 12:356-357.

Davies JL, Kawaguchi Y, Bennett ST, et al. (1994): A genome-wide search for human type 1 diabetes susceptibility genes. Nature 371:130-6.

DeLisi LE, Devoto M, Lofthouse R, et al. (1994): Search for linkage to schizophrenia on the X and Y chromosomes. American Journal of Medical Genetics 54:113-21.

DeLisi LE, Goldin LR, Maxwell ME, et al. (1987): Clinical features of illness in siblings with schizophrenia or schizoaffective disorder. Arch Gen Psychiatry 44:891-6.

American Psychiatric Association (1987): Diagnostic and Statistical Manual (Third Edition - Revised). Washington, D.C.: American Psychiatric Press.

Easton DF, Bishop DT, Ford D, et al. (1993): Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. American Journal of Human Genetics. 52:678-701.

Faraway JJ (1993): Distribution of the admixture test for the detection of linkage under heterogeneity. Genetic Epidemiology 10:75-83.

Garner C, Kelly M, Cardon L, et al. (in press): Linkage analyses of schizophrenia to chromosome 6p24-p22: an attempt to replicate. American Journal of Medical Genetics.

Gejman PV, Martinez M, Cao Q, et al. (1993): Linkage analysis of fifty-seven microsatellite loci to bipolar disorder. Neuropsychopharmacology 9:31-40.

Gill M, McGuffin P, Parfitt E, et al. (1993): A linkage study of schizophrenia with DNA markers from the long arm of chromosome 11. Psychological Medicine 23:27-44.

Gill M, Vallada H, Collier D, et al. (1996): A combined analysis of D22S278 marker alleles in affected sib-pairs: support for a susceptibility locus for schizophrenia at chromosome 22q12. American Journal of Medical Genetics 67:40-5.

Gottesman II, Shields J (1967): A polygenic theory of schizophrenia. Proc Natl Acad Sci USA 58:199-205.

Gurling H, Kalsi G, Hui-Sui Chen A, et al. (1995): Schizophrenia susceptibility and chromosome 6p24-22. Nature Genetics 11:234-5.

Gyapay G, Morissette J, Vignal A, et al. (1994): The 1993-94 Genethon human genetic linkage map. Nature Genetics 7(2 Spec No):246-339.

Hauser ER, Boehnke M, Guo S-W, Risch N (1996): Affected sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. Genetic Epidemiology 13:117-138.

Holmans P (1993): Asymptotic properties of affected-sib-pair linkage analysis. American Journal of Human Genetics 52:362-74.

Holmans P, Clayton, D (1995): Efficiency of typing unaffected relatives in an affected sib-pair linkage study with single-locus and multiple tightly linked markers. Am J Hum Genet 57:1221-1232.

Kalsi G, Mankoo BS, Curtis D, et al. (1995): Exclusion of linkage of schizophrenia to the gene for the dopamine D2 receptor (DRD2) and chromosome 11q translocation sites. Psychological Medicine 25:531-7.

Karayiorgou M, Morris MA, Morrow B, et al. (1995): Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. Proc Nat Acad Sci USA 92:7612-6.

Kendler, K.S., O'Neill A, Burke J, et al.. (1996): Irish Study of High-Density Schizophrenia Families: Field Methods and Power to Detect Linkage. American Journal of Medical Genetics 67:179-190.

Kendler, K.S., McGuire M, Gruenberg AM, et al. The Roscommon Family Study. II. The risk of nonschizophrenic nonaffective psychoses in relatives. Archives of General Psychiatry 50:645-52.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996): Parametric and nonparametric linkage analysis: a unified multipoint approach. American Journal of Human Genetics. 58:1347-63, 1996 Jun.

Kruglyak L, Lander ES (1995): Complete multipoint sib-pair analysis of qualitative and quantitative traits. American Journal of Human Genetics 57:439-54.

Lander E, Kruglyak L (1995): Genetic dissection of complex traits: guidelines for interpreting and reporting linkage findings. Nature Genetics 11:241-247.

Lander E, Kruglyak L (1996): Genetic dissection of complex traits. Nature Genetics, 12, 357-358.

Lander ES (1988): Mapping complex genetic traits in humans. In Davies KE (ed), "Genome Analysis, A Practical Approach ." Oxford: IRL Press, pp.171-189.

Lange K, Weeks D, Boehnke M (1988): Programs for pedigree analysis: MENDEL, FISHER, and dGENE. Genetic Epidemiology 5:471-2.

Lathrop GM, Lalouel JM, Julier C, et al. (1984): Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA 81, 3442-3446.

MacLean CJ, Ploughman LM, Diehl SR, Kendler KS (1992): A new test for linkage in the presence of locus heterogeneity. American Journal of Human Genetics 50:1259-66.

Maier W, Lichtermann D, Minges J, et al. (1993): Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. Archives of General Psychiatry 50:871-83.

Moises HW, Yang L, Kristbjarnarson H, et al. (1995): An international two-stage genome-wide search for schizophrenia susceptibility genes. Nature Genetics 11:321-4.

Mowry BJ, Nancarrow DJ, Lennon DP, et al. (1995): Schizophrenia susceptibility and chromosome 6p24-22. Nature Genetics 11:233-4.

Ott J (1991): "Analysis of human genetic linkage." Baltimore: Johns Hopkins Press.

Pulver AE, Karayiorgou M, Wolyniec PS, et al. (1994a): Sequential strategy to identify a susceptibility gene for schizophrenia: report of potential linkage on chromosome 22q12-q13.1: Part 1. American Journal of Medical Genetics 54:36-43.

Pulver AE, Karayiorgou M, Lasseter VK, et al. (1994b): Follow-up of a report of a potential linkage for schizophrenia on chromosome 22q12-q13.1: Part 2. American Journal of Medical Genetics 54:44-50.

Pulver AE, Nestadt G, Goldberg R, et al. (1994c): Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. J Nerv Ment Dis 182(8):476-8.

Pulver AE, Lasseter VK, Kasch L, et al. (1995): Schizophrenia: a genome scan targets chromosomes 3p and 8p as potential sites of susceptibility genes. American Journal of Medical Genetics 60:252-60.

Riley BP, Rajagopalan S, Mogudi-Carter M, et al. (in press): No evidence for linkage of chromosome 6p markers to schizophrenia in southern African Bantu-speaking families. Psychiatric Genetics,.

Schwab SG, Albus M, Hallmayer J, et al. (1995): Evaluation of a susceptibility gene for schizophrenia on chromosome 6p by multipoint affected sib-pair linkage analysis. Nature Genetics 11:325-7.

St.Clair DM, Blackwood DHR, Muir WJ, et al. (1989) No linkage of chromosome 5q11-13 markers to schizophrenia in Scottish families. Nature 339 305-309.

Straub RE, MacLean CJ, O'Neill FA, et al. (1995): A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. Nature Genetics 11:287-93.

Suarez B, Hodge S (1979): A simple method to detect linkage for rare recessive diseases: an application to juvenile diabetes. Clinical Genetics 15:126-136.

Suarez B, Hampe CL, van Eerdewegh P (1995): Problems of replicating linkage claims in psychiatry. In Gershon ES, Cloninger CR (eds). "New Genetic Approaches to Mental Disorders." Washington, D.C.: American Psychiatric Press.

The Huntington's Disease Collaborative Research Group (1993): A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's Disease chromosomes. Cell 72:971-983.

Tsuang MT, Gilbertson MW, Faraone SV (1991): The genetics of schizophrenia: current knowledge and future directions. Schizophrenia Research 4:157-171.

Webb C, Levinson DF, Gladis M (1993): Schizotypal and paranoid personality in the relatives of patients with schizophrenia and affective disorders: A review. Schizophrenia Research 11:81-92.

Witte JS, Elston RC, Schork NJ (1996): Genetic dissection of complex traits. Nature Genetics 12:355-356.